ELSEVIER

Journal of School Psychology 47 (2009) 19–37

Journal of School Psychology

Effects of public feedback during RTI team meetings on teacher implementation integrity and student academic performance

Gary J. Duhon [a,*], Eric M. Mesmer [c], Lezlee Gregerson [a], Joseph C. Witt [b]

[a] *School Psychology Program, Oklahoma State University, 423 Willard Hall, Stillwater, Oklahoma 74075, United States*
[b] *Louisiana State University, United States*
[c] *Radford University, United States*

Received 11 February 2008; received in revised form 25 June 2008; accepted 17 September 2008

Abstract

As a result of the new revision of IDEA (2004), models of early intervention and response-to-intervention (RTI) have received a great deal of attention in the literature. Although various tiered models have been described in detail, one aspect of RTI that has received little empirical attention is the need to ensure integrity of intervention as part of a team decision making process. One method that has good support for improving treatment integrity is performance feedback (PFB); however, the utilization of PFB within a team context has received very little attention in the literature. Experiment 1 evaluated the effect of PFB delivered within the context of an RTI team on treatment integrity after fidelity had fallen to unacceptable levels. Experiment 2 evaluated the effects of PFB on the maintenance of treatment integrity prior to integrity fall off. Results indicate that PFB can be used as an efficient means of improving or maintaining treatment integrity when applied within a team model.

Keywords: RTI; Integrity; Intervention; Teacher; Student

* Corresponding author. Tel.: +1 405 744 9436.
E-mail address: gary.duhon@okstate.edu (G.J. Duhon).
ACTION EDITOR: Brian K. Martens.

0022-4405/$ - see front matter © 2008 Society for the Study of School Psychology. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsp.2008.09.002

Interventions and the outcomes that follow have always played a central role in the remediation of educational difficulties. Recently, interventions have been utilized more and more as assessment devices in the form of pre-referral interventions (McNamara & Hollinger, 2003) and in response-to-intervention (RTI). Consistent with an intervention-based assessment approach, important decisions affecting students are made, in part, based upon an evaluation of the effect of one or more interventions. When evaluating student responsiveness to an intervention, whether in the context of determining need for evaluation or eligibility for special education, it is crucial that valid decisions are made. In evaluating student responsiveness to an intervention, treatment integrity plays an essential role (Perepletchikova & Kazdin, 2005). To ensure that valid decision making will occur requires a demonstration of the functional relationship between student responsiveness and exposure to the intervention. This ultimately requires the presence of treatment integrity data (Gresham, Donald, MacMillan, Beebe-Frankenberger, & Bocian, 2000; Kazdin, 2003; Moncher & Prinz, 1991). Without documentation of integrity, inferences about student responsiveness and, ultimately, the decisions become nothing more than uninformed guesses.

Although treatment integrity data are important for making informed decisions about responsiveness, high levels of treatment integrity are also crucial to ensure a student's right to due process within an RTI framework. Just as utilizing only half of a cognitive battery to evaluate a student's abilities would be a violation of the student's right to a valid assessment, implementation of only half of the intervention would also result in this same end within the context of an RTI evaluation. In order to ensure accountability within an RTI assessment process, treatment integrity must be known, and preferably occur at high levels.

Although treatment integrity is an important issue, high levels of integrity are not a guarantee. Several studies have demonstrated that when levels of treatment integrity are directly measured, the degree to which the intervention plan is followed can vary from high to completely absent (e.g., Mortenson & Witt, 1998; Noell et al., 2000; Wickstrom, Jones, & LaFleur, 1998). In one of the first studies to do so, Wickstrom et al. directly measured levels of treatment integrity by teachers implementing interventions in a regular classroom setting. Although teacher self report indicated moderate to high levels of adherence to original intervention plans, observed levels of treatment integrity were virtually nonexistent.

A number of researchers have identified PFB (PFB; Mortenson & Witt, 1998; Noell, Duhon, Gatti, & Connell, 2002; Noell, Witt, Gilbertson, Ranier, & Freeland, 1997) and direct training (Sterling-Turner, Watson, & Moore, 2002) as variables that can improve the integrity of intervention steps by teachers. Sterling-Turner et al. compared direct and indirect forms of training of intervention components with four consultation cases and the resulting effect of this training on treatment integrity. In a multiple-baseline design across subjects, the percent of treatment integrity by the consultee was measured following sessions of indirect training and then sessions of direct training. Results indicated that direct training produced much higher levels of treatment integrity than indirect training.

PFB has also proven to be effective for increasing desired performance (Green & Reid, 1991). Originally, PFB was evaluated in organizational or institutional settings, but more recently has been applied in school settings to increase integrity of intervention implementation by teachers (Mortenson & Witt, 1998). With PFB, information regarding processes and results are provided to promote transfer or maintenance of skills and behavior (Arco, 1991; Fleming & Sulzer-Azaroff, 1989). Information provided through PFB may

include review of implementation or outcome data, praise for accurate implementation, corrective feedback for errors in implementation and/or other comments regarding questions of implementation that may arise (Codding, Feinberg, Dunn, & Pace, 2005). PFB has been utilized in educational settings to improve implementation of academic interventions (Mortenson & Witt, 1998; Noell et al., 2000; Witt, Noell, LaFleur, & Mortenson, 1997), use of contingent praise (Jones, Wickstrom, & Friman, 1997; Martens, Hiralall, & Bradley, 1997), behavior-management interventions in general education (DiGennaro, Martens, & McIntyre, 2005; Noell et al., 2002), and in special education (Codding et al., 2005; DiGennaro, Martens, & Kleinmann, 2007), as well as adherence to a data-based problem-solving model (Burns, Peters, & Noell, 2008). Witt et al. evaluated the effects of providing feedback to teachers on their use of academic interventions within a multiple-baseline design. Four teachers were asked to administer an academic intervention to students identified as having difficulty in a particular subject area. After a significant drop in integrity during baseline, the teachers were individually provided with PFB in the form of a daily graph of performance to that point. Following the initiation of feedback, teachers increased their treatment integrity to acceptable levels. This general result has been replicated with less training and support from the consultant (Noell et al., 1997), on a less intensive weekly schedule (Mortenson & Witt, 1998), and with strictly behavioral concerns (Noell et al., 2002).

Despite the need for high levels of treatment integrity and the effectiveness of PFB at improving treatment integrity, there has been little adoption in practice. This is likely due to both pragmatic and philosophical barriers that may make providing PFB, as it has been described in the literature, less palatable for practitioners. Pragmatically, practitioners utilizing PFB would need to have sufficient time to deliver feedback as well as a structure designed to encourage this delivery. However, much of the research on providing PFB to teachers requires daily direct consultant/teacher contact. This contact was designed to measure levels of treatment integrity during intervention implementation and to provide daily feedback to the teacher when integrity dropped below acceptable levels. It is likely that the average school-based consultant does not have sufficient time or resources to maintain contact on such a schedule. Additionally, individual consultation does not provide an obvious structure for delivering this type of information to teachers. Practitioners must find opportunities to provide feedback to teachers outside of their current schedule of activities, and these opportunities must coincide with teachers' opportunities to receive feedback. The coordination of these activities is not always easy and likely reduces the probability that PFB will be utilized. Current procedures used to provide PFB address the concern of treatment integrity after integrity has fallen off. As important as treatment implementation is, it would seem more ethical to engage in activities that prevent treatment integrity from becoming a problem as opposed to responding only after integrity has declined.

In order to proactively promote treatment integrity and provide practitioners more realistic means of utilizing PFB, new methods of providing feedback should be developed. Methods that reduce time demands provide a structure designed to increase opportunities for delivery of PFB, and focus on maintenance of integrity may have an increased likelihood of adoption by practitioners. In fact, the current requirement for daily contact may not be necessary for producing improvements in treatment integrity. Mortenson and Witt (1998) produced high levels of treatment integrity with only weekly feedback. In this study, levels of treatment integrity were monitored daily through the use of fax transmittals

while feedback about treatment integrity was delivered face-to-face only once per week. The RTI or pre-referral intervention team process might also serve to increase the use of PFB by providing obvious opportunities to deliver feedback as well as encouraging delivery of PFB in a proactive manner. The RTI or pre-referral intervention team process typically utilizes interventions that result from recommendations by a team charged with oversight and decision making (Barnett, Daly, Jones, & Lentz, 2004). These recommendations are developed during team meetings which occur frequently and on a regular basis. These frequent team meetings would easily provide the structure necessary to increases potential opportunities for feedback about treatment implementation to be delivered by the school psychologist. This would also allow for treatment integrity to be targeted from the very beginning of the intervention process as opposed to PFB being delivered only in response to poor integrity.

Although Mortenson and Witt (1998) has demonstrated the effectiveness of PFB delivered weekly in a one-on-one setting, it is unknown if weekly PFB will improve integrity when delivered in a group setting such as an RTI meeting. Additionally, it is unknown if the application of PFB in a group setting (e.g., a multidisciplinary RTI team) can effectively maintain high levels of integrity for extended periods of time. If PFB can be delivered successfully within a group setting, this application may provide a vehicle for increased usage and enhanced application of this proven technique. Therefore the purpose of the present two-experiment study was to determine if treatment integrity could be enhanced through the application of PFB publicly during existing RTI or school-based intervention team meetings. Furthermore, this study sought to evaluate the effect of PFB on treatment integrity when delivered proactively (prior to intervention fall off) as part of an existing RTI or school based intervention team. Experiment 1 extended the work of Mortenson and Witt (1998) by examining the effectiveness of weekly PFB applied by an individual consultant publicly during RTI team meetings once treatment integrity had fallen off. Experiment 2 extended the use of PFB in school setting to prevention of poor integrity by examining the effect of proactive PFB applied by an individual consultant publicly during RTI team meetings with the goal of preventing low levels of treatment integrity.

1. Experiment 1

1.1. Method

1.1.1. Participants and setting

1.1.1.1. Participants. Participants for Experiment 1 were three certified elementary school teachers and four general education students from a small rural, mid-western public school district. Forty-three percent of the student population received lunch free or at a reduced charge based on family income. Aside from Ms. Swain, who was African-American, all participants were European-American. Ms. Blane was a kindergarten teacher with 3 years of experience. She referred two males, Noah and Tony, both 6 years old at the time of the study. These students were referred to the intervention team for consultation and intervention due to academic concerns. Noah was having difficulty with number recognition while Tony was referred for inaccurate letter recognition. Ms. Swain, a third

grade teacher with 12 years of experience, referred Theo, an 8-year-old male. Ms. Hannon, also a third grade teacher with 6 years of experience, referred Mike, a 9-year-old male.

1.1.1.2. RTI team meeting. The school site utilized intervention teams which consisted of the referring general education teacher, a special education teacher, the school administrator, a school psychologist and a school psychology practicum student who was the consultant and who also served as a researcher for these studies. The team typically met once every other week and discussed the disposition of new and ongoing intervention cases and evaluated responsiveness; however, for cases included in this study, team meetings occurred once per week. The team's weekly meeting occurred in the guidance counselor's office, and all intervention and data collection took place in the teachers' general education classrooms. The team model being utilized was consistent with the problem-solving model described by Fuchs, Fuchs, Bahr, and Stecker (1990). Fuchs, Mock, Morgan, and Young (2003) explain that within this model, "solutions to instructional and behavioral problems are induced by evaluating students' responsiveness to a four-stage process comprising problem identification, problem analysis, plan implementation, and problem evaluation" (p. 160).

1.1.2. Dependent variables and interobserver agreement

1.1.2.1. Treatment integrity. The primary dependent measure was the percentage of intervention steps implemented accurately by the teachers. An intervention was developed for each student as a result of the collaboration with the teacher, researcher, and the intervention team using procedures described below. Interventions were designed to be implemented once per day and required between 5 and 15 min to implement. Each of these interventions included a protocol describing in detail the steps required to implement the intervention. The referring teacher was responsible for carrying out the intervention which resulted in permanent products specific to that intervention. These permanent products consisted of byproducts of the active steps of the intervention. For example, the intervention developed for Mike was a timed fluency drill with reward. The products of the intervention implementation were a completed fluency probe along with a goal for that session, a fluency score for that session, and an indication as to whether a reward was given. Similar to previous research in this area (e.g., Mortenson & Witt, 1998; Noell et al., 2002, 1997; Witt et al., 1997), these permanent products were designed to provide a record of the intervention steps followed and an indication as to whether the intervention steps were completed accurately. Percentage of intervention steps implemented was then calculated by dividing the number of steps completed accurately by the total number of steps required by the intervention protocol and multiplying by 100. The resulting percentage was considered the amount of intervention implementation that occurred for that session. The researcher assigned to the case collected intervention products at the end of each day. These researchers were doctoral-level graduate students in their 3rd year of training and had been trained in and had experience in consulting with teachers for behavioral and academic concerns. A second researcher independently scored a random sample of 33% of the intervention products distributed across all participants. Interscorer agreement for plan implementation was calculated by dividing the smaller score for each day by the larger score and multiplying the result by 100. Interscorer agreement was 96% (range, 92 to 100%).

1.1.2.2. Student academic performance. For Experiment 1, student academic performance with the targeted skill was measured. As part of the intervention protocols, baseline levels of performance and student response to their specific intervention were measured. Either through teacher scoring or independent evaluation of the permanent products, student data were calculated on a daily basis. These data were calculated as either accuracy (number of items correct divided by the number of items attempted then multiplied by 100%), or fluency (number of digits correct per min), depending on the skill targeted and the specific intervention designed. Due to the nature of these interventions and a reliance on teacher scoring and/or permanent products to evaluate student academic performance, it was not possible to collect data on student academic performance during sessions without intervention implementation. These sessions produced no student data and are not reported or included in evaluations of student academic performance. A second researcher independently scored a random sample of 33% of student academic performance products across all participants. Interscorer agreement for student academic performance was calculated by dividing the smaller score for each session by the larger score and multiplying by 100. Interscorer agreement was 95% (range 92, to 100%).

1.1.3. Pre-intervention assessment and intervention development

In order to identify and define a target behavior, the referred student's teacher was interviewed to collect background information regarding the referral concern. Once the researcher completed the interview, an unstructured classroom observation was conducted. This allowed the researcher to become familiar with the student and the classroom environment in an effort to assist in the later development of the intervention.

Following the observation, the case was discussed at the next intervention team meeting. At this meeting the teacher, researcher, and remainder of the intervention team utilized teacher and researcher data to develop an individualized student intervention. Once the intervention was developed as part of the intervention team meeting, the teacher and researcher worked independently to build specific intervention components. Because each student presented different concerns, the referral concern and the interventions developed were different for each student. A summary of the referral concerns and the interventions developed for each participating student are described below and summarized in Table 1.

Table 1
Experiment 1 interventions and outcomes

Teacher	Student	Referral	Intervention	Baseline	Last 3 int. sessions	Increase
Blane	Noah	Number recognition	Flash card drill w/ reward	69% accuracy	92% accuracy	33%
Blane	Tony	Letter recognition	Flash card drill w/ reward	68% accuracy	100% accuracy	47%
Swain	Theo	Accuracy of seat work	Repeated directions, feedback and reward	57% accuracy	80% accuracy	40%
Hannon	Mike	Math fluency	Timed fluency drills w/ reward	9 dcpm/4 errors	16 dcpm/2 errors	78%

1.1.3.1. Experiment 1 interventions. Noah's target behavior concern was accuracy of number recognition to 30, while Tony's was accuracy of letter recognition (both upper case and lower case). The intervention team met to discuss Noah's and Tony's concerns separately. As a result of teacher and consultant data, it was decided that although the academic problems were in two different domains, the deficits were similar enough to allow for similar interventions for both students. This would also assist the teacher in administration because only one protocol would have to be learned. For both cases a technique called drill sandwich (MacQuarrie, Tucker, Burns, & Hartman, 2002) was used in which unknown digits or letters were folded into a flashcard drill procedure. Both students began with 10 flash cards, which included 7 knowns and 3 unknowns based upon a daily assessment. Using the drill sandwich technique, the students were presented with the flash cards repeatedly until they could accurately identify all 10 letters or numbers. New unknowns were added each session and old knowns were removed. Prior to each daily session, an evaluation of all items to be learned was collected. Ms. Blane was provided with an intervention folder for each student which contained the flashcards and a protocol with 10 steps for administration and scoring of the intervention product. A consultant collected the intervention protocols and the student products after each daily session.

The target behavior for Theo was accuracy of academic work when completing language-arts assignments. Based on the results of the pre-intervention assessment activities, it was discovered that Theo was capable of completing language-arts seatwork assignments accurately; however, poor attentiveness to instructions and insufficient feedback about performance resulted in Theo completing his assignments inaccurately. The team developed an intervention that focused on first ensuring Theo understood the assignment instructions by checking for understanding and, second, increasing the amount of feedback about performance provided by Ms. Swain during the intervention. Feedback was provided in the form of rewards when Theo met an accuracy criterion of 80% for accuracy and simple correction when performance dropped below the established criterion. The intervention took place daily during language-arts periods when independent seatwork was assigned, and afterwards Ms. Swain placed the competed language-arts assignment along with an eight-step protocol check sheet in an intervention folder that was checked daily by the consultant.

Mike's target behavior was mathematics fluency, specifically subtraction with digits up to 18. Although the result of the pre-intervention assessment activities identified several areas of math performance deficits, subtraction with digits up to 18 was identified by the teacher as the area of greatest concern. As a result, the intervention team chose a fluency-building intervention designed to provide Mike with more opportunities to practice this already accurate skill. The intervention consisted of providing Mike with a math probe with subtraction problems from 18 and instructing him that he had 2 min to complete as many problems as possible. A goal for performance was also provided which specified both accuracy and fluency requirement for receiving a reward upon completion of the 2 min time period. Ms. Hannon was provided with an intervention folder which contained the math probes and a protocol with 10 steps for administration and scoring of the intervention product. A consultant collected the intervention protocols and the student products after each daily session.

1.1.3.2. Teacher training. Prior to initial implementation and one day prior to the next intervention team meeting, an in-class training day was conducted with each teacher. The training was conducted to ensure that all involved personnel understood their role in the intervention and any minor "on the spot" modifications could be made to the intervention. Prior to the training day, the researcher provided the teacher with all necessary materials to complete the intervention, including teacher implementation protocols, student daily performance report forms, intervention worksheets or flash cards, and reinforcers as necessary and then reviewed the intervention steps with the teacher and the student. On the first day of implementation, the researcher visited the classroom when the intervention was scheduled to occur to help the teacher implement the plan with complete accuracy. If the teacher did not implement any element of the intervention, the researcher immediately reminded the teacher of the treatment step and asked the teacher to complete that intervention step.

1.1.3.3. Intervention implementation. At the next intervention team meeting, the participating teachers were asked to use the materials provided to implement the intervention independently. Teachers were also asked to attend future intervention team meetings weekly in order to review student response to intervention.

1.1.4. Experimental design and procedure

The effect of PFB on treatment integrity delivered in the context of the intervention team was evaluated with a multiple-baseline design across subjects (Barlow & Hersen, 1984). As with most of the previous research in this area, the multiple-baseline design allowed for evaluation of the effect of PFB on treatment integrity after integrity had dropped off to unacceptable levels during an independent implementation phase. Once a treatment effect was established, PFB was removed to evaluate maintenance of the treatment effect in the absence of PFB.

1.1.4.1. Independent implementation phase. During independent implementation, intervention implementation and student outcome data were collected by a researcher daily. No other contact with the teacher occurred during baseline for participants in Experiment 1.

1.1.4.2. Performance feedback. If treatment integrity became variable or fell below acceptable levels (i.e., below 70% integrity for at least 2 consecutive days), PFB was initiated within the constraints of the experimental design. Once integrity fell to unacceptable levels, PFB was delivered at the very next intervention team meeting (within 1–5 days) by the researcher who presented two graphs. The first graph depicted student performance on the target behavior described above (see individual student interventions) in relation to baseline levels of target behavior. The second graph depicted the percentage of treatment steps implemented by the teacher each day. The percentage of treatment steps implemented was calculated as described above (see dependent variables). The researcher presented the teacher and each team member with both graphs and indicated that integrity was below acceptable levels. The researcher then identified the treatment steps most often missed or completed incorrectly over the preceding implementation period and discussed

with the teacher and the team how implementation might be improved for the week ahead. Once initiated, PFB at subsequent intervention team meetings consisted of the same two graphs described above updated with the current data and reviewed in the same manner. If implementation was at acceptable levels, the teacher was thanked by the researcher for accurate implementation and the student's performance became the focus of the discussion.

1.1.4.3. Feedback removal. The delivery of PFB was discontinued once treatment integrity was at or above acceptable levels. At the next possible intervention team meeting, three of the four teachers were asked to maintain implementation of the intervention. Ms. Swain was not asked to maintain the intervention because Theo's work completion was considered to be at criterion levels as determined by the team and intervention was no longer necessary. The remaining teachers were told that they would no longer need to attend intervention team meetings to review the data unless they believed that the student could benefit from such a review. However, Ms. Blane and Ms. Hannon were informed that the treatment integrity data and student data would continue to be collected by the researcher. Teacher treatment integrity data and student outcome data were still collected daily for these teachers and no other contact was made with the teacher.

1.2. Results

1.2.1. Intervention implementation

Results of teacher treatment integrity for Experiment 1 are displayed in Fig. 1. During the first session of teacher training, all teachers implemented all elements of the intervention with 100% integrity and subsequently began independent implementation baseline. Ms. Blane's implementation of Noah's intervention during baseline only occurred on the third day, with 90% integrity, and no implementation occurred during the remaining baseline sessions (M=18%, SD=40). Following the initiation of PFB at the next intervention team meeting, an immediate increase in implementation was observed and sustained during this phase (M=96%, SD=8.4). When feedback was removed, implementation dropped off immediately and the intervention was no longer conducted.

Ms. Blane's initial implementation of Tony's intervention occurred on the first and third day of baseline with 60% and 70% integrity, respectively, and no implementation for the remaining baseline sessions (M=18%, SD=31). As a result of the application of PFB at the next intervention team meeting, an immediate increase in implementation was observed (M=100%, SD=0). As with Ms. Blane's implementation of the previous intervention, when feedback was removed implementation dropped off immediately and the intervention was no longer conducted.

Ms. Swain's implementation was initially at high levels for the first six sessions, but dropped off to 0 for the next six sessions (M=47%, SD=50). At the next intervention team meeting, PFB was initiated, which was followed by an increase in implementation (M=100%, SD=0). Three data points in this phase are missing because Theo was absent due to illness.

Ms. Hannon's implementation was initially variable, and after several sessions of independent implementation dropped off completely (M=26%, SD=33). PFB was initiated and resulted in increases in treatment integrity; however, initial response to feedback was

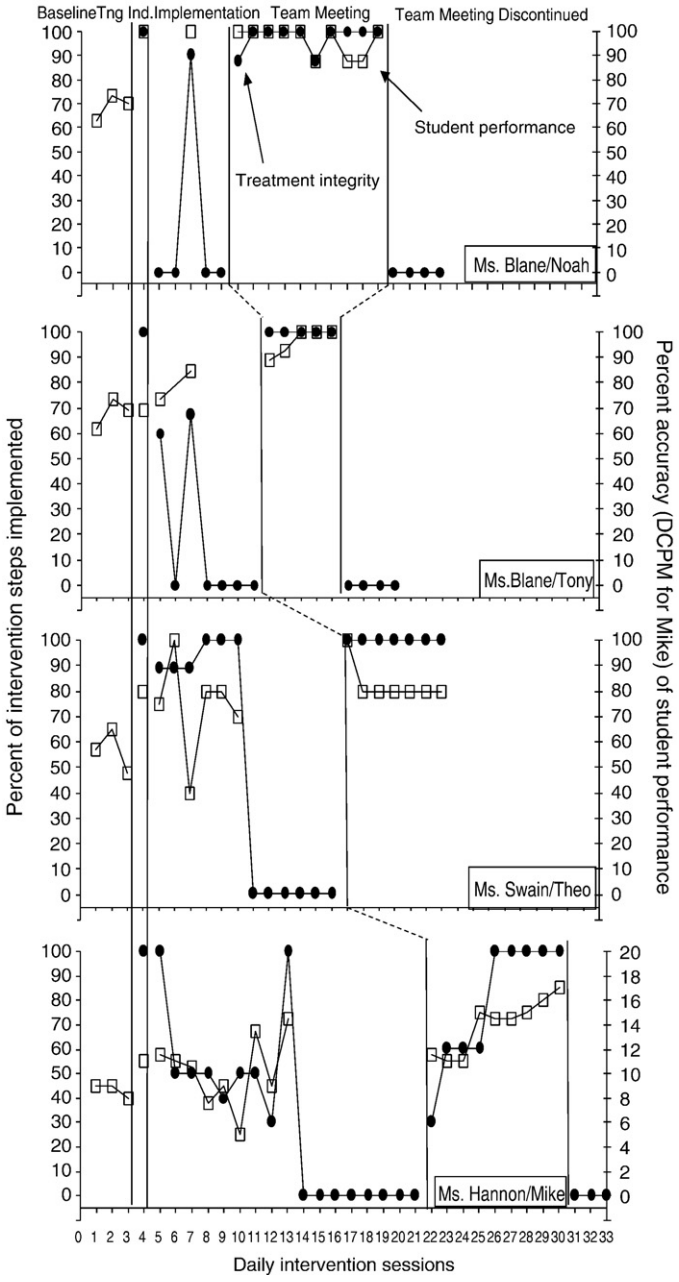

Fig. 1. Percent of integrity of intervention implementation for teachers in Experiment 1.

still below satisfactory levels. Following the next session of PFB, implementation improved to acceptable levels, with overall mean of 79% (SD=26). Implementation of the final five sessions was much higher (M=100%; SD=0). As with previous cases, once PFB was removed the intervention was no longer conducted.

1.2.2. Student academic performance

Baseline levels of student academic performance and performance during the last 3 sessions of PFB are reported in Table 1. A comparison of baseline performance levels to that of the final implementation sessions demonstrates an improvement in the dependent measure for all students in Experiment 1. Because student data were collected throughout the entire study, daily measures of student performance are also presented in Fig. 1.

2. Experiment 2

Experiment 1 examined the impact of PFB on treatment integrity once unacceptable levels were observed. Experiment 2 attempted to examine the effect of proactive PFB applied during an RTI team meeting prior to integrity fall off with the goal of preventing low levels of treatment integrity.

2.1. Method

2.1.1. Participants and setting

2.1.1.1. Participants. Participants for Experiment 2 were four certified elementary school teachers and four general education students from the same small rural, mid-western public school district described above. All participants, both teachers and students, were European-American. Ms. Prat was a first grade teacher with 13 years of experience. She referred Terry, a 6-year-old male. Ms. Mills, a third grade teacher with 7 years of experience referred Dallas, an 8-year-old male. Ms. Jones, a first grade teacher who had 5 year experience referred Cody, a 7-year-old male. Ms. Warner, a first grade teacher with 15 years experience referred Tom, a 6-year-old male. These students were also referred to the intervention team for consultation and intervention due to academic concerns. During an initial interview, participating teachers in both Experiment 1 and Experiment 2 described these problems as being severe enough to warrant individualized intervention.

2.1.1.2. RTI team meetings. The RTI team meetings in Experiment 2 were identical to the RTI team meetings described in Experiment 1.

2.1.2. Dependent variables and interobserver agreement

2.1.2.1. Treatment integrity. The dependent variable of treatment integrity was measured in the same manner as in Experiment 1. A second researcher independently scored a random sample of 33% of the intervention products distributed across all participants. Scorer agreement for plan implementation was calculated by dividing the smaller score for each day by the larger score and multiplying the result by 100. Interscorer agreement was 99% (range, 97 to 100%).

2.1.2.2. Student academic performance. The dependent variable of student academic performance was measured in the same manner as in Experiment 1. A second researcher independently scored a random sample of 33% of student academic performance products across all participants. Interscorer agreement for student academic performance was calculated by dividing the smaller score for each session by the larger score and multiplying the result by 100. Interscorer agreement was 97% (range, 95 to 100).

2.1.3. Pre-intervention assessment and intervention development

Pre-intervention assessment and intervention development procedures in Experiment 2 were identical to those in Experiment 1. Because each student presented different concerns, the referral concerns and the interventions developed were different for each student. A summary of the referral concerns and the interventions developed for each participating student are described below and summarized in Table 2.

2.1.3.1. Experiment 2 interventions. The target behavior for Terry and Tom was the percentage of work completed accurately during independent seat work activities. Based on the results of the pre-intervention assessment activities, it was determined that both students were capable of completing independent seat work within the time allowed and with acceptable accuracy when provided with more frequent monitoring from the teacher and rewarded for increased on-task behavior and accuracy. The team developed interventions for both students that focused on structuring more frequent interactions between the teachers and their target students in order to monitor and reward increased amounts of work completed accurately. A card mounted on each student's desk was used to structure the monitoring intervals and provide a product to track teacher implementation as well as student outcomes. The consultant removed and scored the card containing the 18 steps of the intervention at the end of each day as a measure of treatment integrity and student performance.

Dallas's target behavior was accuracy of telling time on a teacher administered seat-work assignment. Based on the results of the pre-intervention assessment, it was determined that Dallas did not have the skill necessary to tell time on an analog clock. The team developed a 12-step intervention that provided Dallas with instruction designed to provide the skill necessary to tell time. Through the use of modeling and feedback, Ms. Mills instructed Dallas on telling time to the minute given a model of a clock. A folder was developed that contained several weeks of the needed worksheets and daily teacher

Table 2
Experiment 2 interventions and outcomes

Teacher	Student	Referral	Intervention	Baseline	Last 3 int. sessions	Increase
Prat	Terry	Accuracy of seat work	Goal setting w/ reward	43% accuracy	83% accuracy	93%
Mills	Dallas	Reading analog clock	Modeling and feedback	7% accuracy	80% accuracy	1043%
Jones	Cody	Letter writing accuracy	Practice, corrective feedback w/ reward	57% accuracy	89% accuracy	56%
Warner	Tom	Seat work accuracy	Goal setting w/ reward	52% accuracy	90% accuracy	73%

protocols that were collected daily by the consultant to score both integrity and student performance.

Cody's target behavior was accuracy of letters written during a daily written language assignment. In reviewing pre-intervention assessment results, the intervention team selected intervention strategies which would provide the student more opportunities for practice with corrective feedback and reward for improved performance. The 16-step intervention consisted of reviewing Cody's daily written language assignment after it was complete and providing feedback about letters formed incorrectly and an opportunity to rewrite the incorrectly formed letters. A goal for percent of accuracy was established by the teacher and if exceeded Cody would receive a reward. Intervention protocols as well as the permanent product from the intervention were placed in a specified folder and collected daily by the consultant.

2.1.3.2. Teacher training. As in Experiment 1, prior to initial implementation and one day prior to the next intervention team meeting, an in-class training day was conducted with each teacher. All other training activities occurred in the same manner as Experiment 1.

2.1.3.3. Intervention implementation. Intervention implementation was initiated at the next intervention team meeting when the participating teachers were asked to use the materials provided to implement the intervention independently. Teachers were also asked to attend future intervention team meeting weekly in order to review student response to intervention.

2.1.4. Experimental design and procedure

The effect of PFB, on treatment integrity, delivered in the context of the intervention team was evaluated with a multiple-baseline design across subjects (Barlow & Hersen, 1984). However, in Experiment 2, because the intent was to examine the contribution of PFB to maintenance of treatment integrity prior to any initial implementation drop off, PFB was applied prior to an independent implementation phase. As in Experiment 1, once treatment integrity remained at stable levels in the presence of PFB, the feedback was removed and the effect on treatment integrity was evaluated.

2.1.4.1. Performance feedback. At the first meeting following initial implementation (ranged between 1 and 4 days of implementation), teachers in Experiment 2 were provided PFB in an attempt to assure high treatment integrity prior to any intervention fall off. PFB was delivered in the same manner as in Experiment 1 and at the same intervals.

2.1.4.2. Feedback removal. For teachers who did not fall below the 70% criterion for more than 2 consecutive days within the multiple-baseline design, feedback was removed. As in Experiment 1, teachers were asked to maintain the intervention and told that integrity data would still be collected, but that they would no longer be asked to review the data in intervention team meetings unless they felt the case could benefit from such a review. Teacher treatment integrity data and student outcome data were still collected at the end of the day, no other contact was made with the teacher, and the data were not available to anyone but the researcher.

2.1.4.3. Integrity of experimental procedures. Integrity checks were performed on the experimental procedures for data review and PFB. For both procedures, a checklist was used and an independent observer scored whether or not the researcher performed each step in the procedure. Integrity of implementation was assessed for each meeting across all cases. All researchers completed 100% of the items on the integrity checklist.

2.2. Results

2.2.1. Intervention implementation

Results for Experiment 2 are displayed in Fig. 2. All teachers implemented all elements of the intervention during the first or second session of teacher training and subsequently began implementing the intervention in the PFB condition. In this condition Ms. Prat's implementation initially fell off to 72%, but following the first intervention team meeting

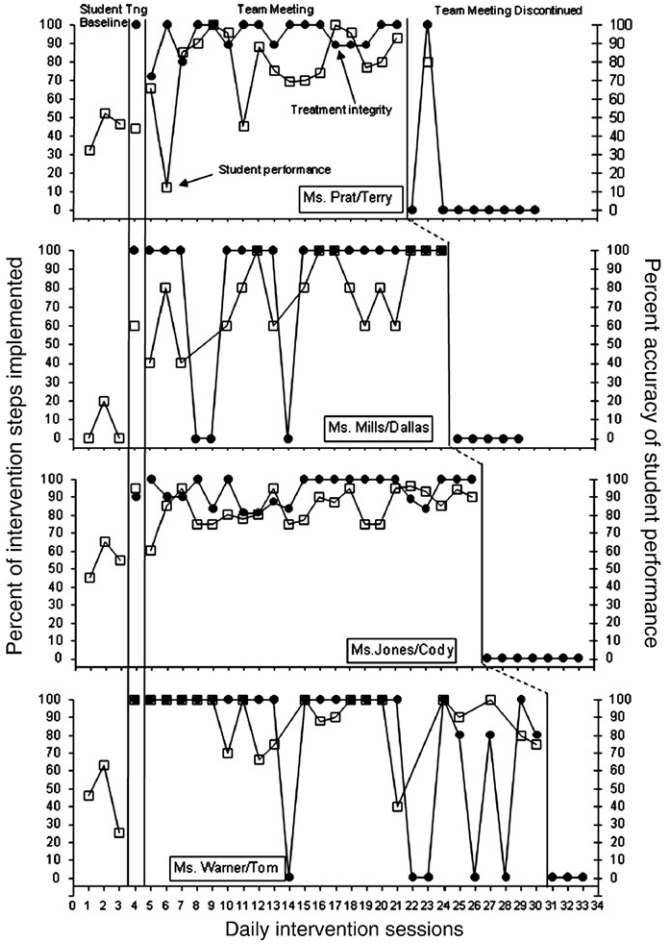

Fig. 2. Percent of integrity of intervention implementation for teachers in Experiment 2.

integrity increased to high and only slightly variable levels (M=94%; SD=8.5). After 4 intervention team meetings the PFB was removed and intervention implementation dropped off to 0 except for one session of implementation at 100%.

Ms. Mills implemented the intervention with 100% integrity throughout the feedback phase except for 3 days where 0 were recorded, reflecting no implementation on those days. Overall implementation was high (M=84%; SD=37) until PFB was removed after the 4th intervention team meeting resulting in a complete fall off in implementation.

Ms. Jones implemented the intervention with high levels of treatment integrity throughout the feedback condition (M=94%; SD=8). Feedback was removed after 5 intervention team meetings and subsequently implementation immediately fell off to 0.

Ms. Warner implemented the intervention with high levels of integrity, but as this phase grew in length, variability began to increase (M=78%; SD=39). Once PFB was removed following the 6th intervention team meeting, implementation fell off completely and the intervention was no longer run.

2.2.2. Student academic performance

Baseline levels of student academic performance and performance from the last 3 sessions with acceptable levels of integrity are reported in Table 2. A comparison of baseline levels to that of the final implementation sessions demonstrates an improvement in the dependent measure for all students in Experiment 2. Because student data were collected throughout the entire study, daily measures of student performance are also presented in Fig. 2.

3. General discussion

This study attempted to replicate and extend previous research on consultation and intervention implementation by utilizing PFB in the context of an RTI team process for the purpose of improving or maintaining high levels of teacher treatment integrity. Similar to previous work, teachers in Experiment 1 failed to implement interventions accurately for extended periods of time (Mortenson & Witt, 1998; Noell et al., 1997, 2000; Witt et al., 1997). After high levels of intervention implementation during the training phase, Ms. Blane and Ms. Hannon both dropped off to unacceptable levels during baseline, with Ms. Swain implementing at high levels for six sessions before the intervention was no longer run. Also consistent with previous research, treatment integrity improved immediately to acceptable levels for all but Ms. Hannon following the application of PFB. Ms. Hannon, however, did improve to acceptable levels after a second intervention team meeting with PFB. Results of the effect of PFB on teacher integrity for Experiment 1 strengthen support for the use of weekly PFB for improving teacher treatment integrity.

Results from Experiment 1 also expand previous research by applying weekly PFB publicly during RTI team meetings after initial fall off. With previous research, PFB was delivered by a single individual in a one-to-one context. Experiment 1 data suggest that the context of an RTI team can serve as an effective delivery system for PFB when improving poor implementation of an intervention is the goal. Delivering PFB publicly during the RTI team meetings cannot only improve integrity but may benefit consultation in other ways. First, the structure of the RTI team meeting once a week is logistically more feasible than

attempting to meet on a daily basis to provide PFB. As a result of this predictable exchange, more frequent and better communication between service delivery providers may occur. Second, the team meeting created a level of public accountability for intervention implementation. Administration, colleagues and other professionals were all made aware of low (and high) levels of implementation, which may have produced pressure to implement the interventions with high degrees of integrity.

Experiment 2 results extend previous research, as well as extending the results of Experiment 1, by successfully applying PFB in a design which maintained intervention implementation for an extended period. For teachers involved in Experiment 2, intervention implementation occurred well beyond what typically has occurred in previous work on treatment integrity (e.g. Mortenson & Witt, 1998; Noell et al., 2002, 2000, 1997; Witt et al., 1997). Combined with the almost immediate drop-off of implementation after removal of feedback for seven of the cases (across both Experiment 1 and Experiment 2) this result suggests that PFB delivered in the context of an RTI team can effectively increase or maintain integrity either after integrity has become a problem or prior to implementation fall-off.

Although PFB delivered in the context described above resulted in outcomes similar to that of previous work, one different outcome was the immediate drop off that occurred once feedback was removed. This is significant in light of the fact that an effective intervention was in place for seven students. Two factors may have contributed to this outcome. First, the delivery of PFB was associated with very salient stimuli (i.e., RTI team meeting), and the removal of PFB was an equally salient change in the stimulus environment. These changes would have made discrimination very easy for the teachers and this discrimination may have resulted in the virtually immediate change in behavior that occurred. Second, many of the students' original target behaviors were dramatically improved as a result of this process. Once the teacher concerns were minimized, the teacher may have seen the intervention as unnecessary. If problem severity is related to high degrees of integrity as suggested by Elliott, Turco, and Gresham (1987), then it stands to reason that the lack of a problem may have the opposite effect.

Although these results are very promising, several limitations and areas of future research remain. The use of multiple-baselines across subjects provides support for internal validity; however, the external validity of these results may be limited. Nonetheless, convergence with previous research regarding the use of PFB to improve treatment integrity supports the contention that PFB may be useful with a variety of populations and problem types. Further, results from Experiment 1 and Experiment 2 expand the means by which PFB may be delivered.

A second limitation is the reliance on permanent products, and specifically teacher reported data to collect the dependent measures. Although permanent products have been utilized for several similar studies in the past (e.g. Mortenson & Witt, 1998; Noell et al., 2002, 1997; Witt et al., 1997), they are still indirect measures of behavior because they do not involve the direct observation of behavior.

A third limitation of this study is that during the feedback removal phase for both Experiment 1 and Experiment 2, feedback and required attendance at the team meetings were removed simultaneously. Because of this methodological concern it is impossible to draw conclusions regarding the variable effecting change in teacher behavior during this phase. However, because the focus of this study was on improving or maintaining teacher

integrity, this methodological concern is limited to the feedback removal phase and is not central to the primary purpose of this study or the interpretation of the results.

A final limitation of this study is that although the goal was to increase treatment integrity, results do not improve our understanding of what the necessary or minimum levels are of treatment integrity required to improve student achievement. Although it would be logical to assume that treatment integrity is related to student academic performance, the design utilized and the method of measuring student academic performance for Experiment 1 and Experiment 2 do not allow for an evaluation of this relationship. Treatment integrity was not an independent variable manipulated within an experimental design and student academic performance was only measured in the presence of measurable levels of treatment integrity. This lack of control with regard to the measurement of student academic performance makes it impossible to demonstrate a true causal relationship between the two variables.

Future research should be conducted to assist in identifying the mechanism underlying PFB. Whether the delivery of PFB serves as a reward, punisher, or if the effect of PFB is specific to individuals is currently unknown. PFB is a general procedure that may serve to affect integrity differently depending upon the particular situation in which it is applied. For example, if a teacher is unaware of an error in implementation, PFB may serve to make them aware of this error and improved implementation may result. This would be consistent with a skill-based hypothesis of PFB effectiveness. However, it may be that the delivery of PFB is an aversive stimulus and the removal of this stimulus may serve as a negative reinforcer for implementation. This would be consistent with a performance-based hypothesis of PFB effectiveness. DiGennaro et al. (2005, 2007) have attempted to evaluate this very hypothesis by examining the effects of PFB delivered alone and in conjunction with a negative reinforcer. Results from this work indicate that PFB alone improved integrity minimally for some teachers, but PFB combined with the negative reinforcer of repeated practice improved integrity dramatically for all teachers. Understanding the mechanisms underlying PFB's effectiveness will assist in improving these techniques and providing more options for researchers and practitioners alike. For instance, if the underlying mechanism of PFB is negative reinforcement then variable or random schedules could be utilized to maintain high levels of integrity for extended periods of time with minimal effort. More research in this area is crucial for improving the effectiveness of PFB.

The effectiveness of PFB in the context of an intervention team was supported within this work; however, it is unknown what effect this delivery will have on future referrals to the team or to future intervention implementation. The public interaction may result in a reduced desire to interact with the team or researcher in the future. The end result would be that students with difficulties may not be referred for assistance as a result of a bias against the team or researcher. On the other hand, the use of PFB in a proactive and preventative manner may actually increase teachers' perceptions of positive support, thereby increasing their use of intervention services.

Finally, these results are particularly important in light of the emphasis on intervention and, particularly, RTI. To make evaluations about responsiveness, not only must the levels of implementation be known, but the decision makers must be confident that the interventions were implemented as planned. The use of PFB has been one of only a few methods empirically validated for achieving consistently high levels of integrity, and its

expanded use can only assist in improving RTI and ultimately services for students in need. For example, Burns et al. (2008) successfully applied PFB to increase adherence to the problem-solving model. If PFB procedures described within the current study were combined with those from Burns et al., it could potentially result in high levels of integrity for both RTI model adherence and for the interventions that are contained within. As with any form of decision making, for RTI decisions to be valid they must first be reliable and to achieve reliability integrity is a must. In the end, the use of PFB procedures may serve as procedural safeguards and enhance reliability and validity of RTI.

References

Arco, L. (1991). Effects of outcome performance feedback on maintenance of client and staff behavior in a residential setting. *Behavioral Residential Treatment*, *6*, 231–247.

Barlow, D. H., & Hersen, M. (1984). *Single case experimental designs: Strategies for studying behavior change.* New York: Pergamon Press.

Barnett, D. W., Daly, E. J., III, Jones, K. M., & Lentz, F. E., Jr. (2004). Response to intervention: Empirically based special service decisions from single-case designs of increasing and decreasing intensity. *Journal of Special Education*, *38*, 66–79.

Burns, M. K., Peters, R., & Noell, G. H. (2008). Using performance feedback to enhance implementation fidelity of the problem-solving team process. *Journal of School Psychology*, *46*, 537–550.

Codding, R. S., Feinberg, A. B., Dunn, E. K., & Pace, G. M. (2005). Effects of immediate performance feedback on implementation of behavior support plans. *Journal of Applied Behavior Analysis*, *38*, 205–219.

DiGennaro, F. D., Martens, B. K., & Kleinmann, A. E. (2007). A comparison of performance feedback procedures on teachers' treatment implementation integrity and students' inappropriate behavior in special education classrooms. *Journal of Applied Behavior Analysis*, *40*, 447–461.

DiGennaro, F. D., Martens, B. K., & McIntyre, L. L. (2005). Increasing treatment integrity through negative reinforcement: Effects on teacher and student behavior. *School Psychology Review*, *34*, 220–231.

Elliott, S. N., Turco, T. L., & Gresham, F. M. (1987). Consumers' and clients' pretreatment acceptability ratings of classroom group contingencies. *Journal of School Psychology*, *25*, 145–153.

Fleming, R. K., & Sulzer-Azaroff, B. (1989). Enhancing quality of teaching by direct care staff through performance feedback on the job. *Behavioral Residential Treatment*, *4*, 377–395.

Fuchs, D., Fuchs, L. S., Bahr, M. W., & Stecker, P. M. (1990). Prereferral intervention: A prescriptive approach. *Exceptional Children*, *56*, 493–513.

Fuchs, D., Mock, D., Morgan, P. L., & Young, C. L. (2003). Responsiveness-to-intervention: Definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities Research and Practice*, *18*, 157–171.

Green, C. W., & Reid, D. H. (1991). Reinforcing staff performance in residential facilities: A survey of common managerial practices. *Mental Retardation*, *29*, 195–200.

Gresham, F. M., Donald, L., MacMillan, D. L., Beebe-Frankenberger, M. E., & Bocian, K. M. (2000). Treatment integrity in learning disabilities intervention research: Do we really know how treatments are implemented? *Learning Disabilities Research and Practice*, *15*, 198–205.

Jones, K. M., Wickstrom, K. F., & Friman, P. C. (1997). The effects of observational feedback on treatment integrity in school-based behavioral consultation. *School Psychology Quarterly*, *12*, 316–326.

Kazdin, A. E. (2003). *Research design in clinical psychology 4th ed.).* Boston, MA: Allyn & Bacon.

MacQuarrie, L. L., Tucker, J. A., Burns, M. K., & Hartman, B. (2002). Comparison of retention rates using traditional drill sandwich, and incremental rehearsal flash card methods. *School Psychology Review*, *31*, 584–595.

Martens, B. K., Hiralall, A. S., & Bradley, T. A. (1997). A note to teacher: Improving student behavior through goal setting and feedback. *School Psychology Quarterly*, *12*, 33–41.

McNamara, K., & Hollinger, C. (2003). Intervention-based assessment: Evaluation rates and eligibility findings. *Exceptional Children*, *69*, 181–193.

Moncher, F. J., & Prinz, R. J. (1991). Treatment fidelity in outcome studies. *Clinical Psychology Review*, *11*, 247–266.

Mortenson, B. P., & Witt, J. C. (1998). The use of weekly performance feedback to increase teacher implementation of a prereferral academic intervention. *School Psychology Review*, *27*, 613−627.

Noell, G. H., Duhon, G. J., Gatti, S. L., & Connell, J. E. (2002). Consultation, follow-up, and implementation of behavior management interventions in general education. *School Psychology Review*, *31*(2), 217−234.

Noell, G. H., Witt, J. C., Gilbertson, D. N., Ranier, D. D., & Freeland, J. (1997). Increasing teacher intervention implementation in general education settings through consultation and performance feedback. *School Psychology Quarterly*, *12*, 77−88.

Noell, G. H., Witt, J. C., LaFleur, L. H., Mortenson, B. P., Rainer, D. D., & LeVelle, J. (2000). Increasing intervention implementation in general education following consultation: A comparison of two follow-up strategies. *Journal of Applied Behavior Analysis*, *33*, 271−284.

Perepletchikova, F., & Kazdin, A. E. (2005). Treatment integrity and therapeutic change: Issues and research recommendations. *Clinical Psychology: Science and Practice*, *12*, 365−383.

Sterling-Turner, H. E., Watson, S. T., & Moore, J. W. (2002). The effects of direct training and treatment integrity on treatment outcomes in school consultation. *School Psychology Quarterly*, *17*, 47−77.

Wickstrom, K. F., Jones, K. M., & LaFleur, L. H. (1998). An analysis of treatment integrity in school-based behavioral consultation. *School Psychology Quarterly*, *13*, 141−154.

Witt, J. C., Noell, G. H., LaFleur, L. H., & Mortenson, B. P. (1997). Teacher use of interventions in general education settings: Measurement and analysis of the independent variable. *Journal of Applied Behavior Analysis*, *30*, 693−696.